B.S.T.J. BRIEF

Solving Nonlinear Network Equations Using Optimization Techniques

By ALLEN GERSHO

(Manuscript received September 10, 1969)

A class of nonlinear equations arising in transistor network analysis, as well as in other areas, has the form

$$f_i(x_i) + \sum_{j=1}^{n} a_{ij}x_j - b_i = 0 \qquad i = 1, 2, \cdots, n \tag{1}$$

or in matrix notation

$$\mathbf{F}(\mathbf{x}) + A\mathbf{x} - \mathbf{b} = \mathbf{0}, \tag{2}$$

where the nonlinearities $f_i(\cdot)$ are continuously differentiable, strictly monotone increasing functions. Results by Willson[1] and Sandberg and Willson[2,3] on nonlinear networks have included broad conditions for the existence and uniqueness of a solution to equation (2). However, convergent computational algorithms for finding the solution have been given only for restricted subclasses of the class of equations that have unique solutions.[1,2,4,5] These subclasses are characterized by a variety of restrictions on the matrix A and on the type of nonlinearities. In this brief we show that a single convergent algorithm exists for solving these equations under conditions virtually as broad as the known existence and uniqueness conditions. Peripherally, we obtain under these conditions a conceptually simple proof of the existence of a solution.

The approach is to use the old technique (probably due to Cauchy) of converting a root-finding problem to a minimization problem. Let

$$\mathbf{r}(\mathbf{x}) \triangleq \mathbf{F}(\mathbf{x}) + A\mathbf{x} - \mathbf{b}, \tag{3}$$

and define the scalar valued "potential" function

$$Q(\mathbf{x}) \triangleq \mathbf{r}^T B \mathbf{r} \tag{4}$$

where B is an arbitrarily chosen symmetric positive definite matrix and T denotes the transpose. Then $Q(\mathbf{x})$ is positive unless x is a solution of equation (2). Consequently, minimizing $Q(\mathbf{x})$ is equivalent to solving equation (2) if in fact the nonlinear equation (2) has a solution.

Since $Q(\mathbf{x})$ is continuous, we may regard it as a continuous surface and observe that if

$$Q(\mathbf{x}) \to \infty \quad \text{as} \quad ||\mathbf{x}|| \to \infty \tag{5}$$

the so-called "level sets",

$$\{\mathbf{x} : Q(\mathbf{x}) < c\},$$

are bounded for each number $c > 0$ and there must exist a point $\mathbf{x}^*$ where $Q(\mathbf{x})$ attains a global minimum. Under what conditions will this minimum satisfy $Q(\mathbf{x}^*) = 0$ so that $\mathbf{x}^*$ is a solution of equation (2)? From equations (3) and (4) the gradient of Q is easily found to be

$$\nabla Q(\mathbf{x}) = 2(D_{\mathbf{x}} + A^T)B\mathbf{r} \tag{6}$$

where $D_{\mathbf{x}}$ is the positive diagonal matrix whose ith diagonal element is $f_i'(x_i)$ where the prime denotes differentiation. Since the gradient must be zero at a minimum, either (i)

$$\mathbf{r}(\mathbf{x}^*) = \mathbf{0},$$

or (ii)

$$\det\{D_{\mathbf{x}} + A\} = 0 \quad \text{at} \quad \mathbf{x} = \mathbf{x}^*.$$

If A is in the class of matrices P_0 characterized by the property[3]

$$\det\{D + A\} \neq 0 \text{ for all diagonal matrices } D > 0, \tag{7}$$

it follows that condition (i) holds so that $\mathbf{x}^*$ is a solution of equation (3) for A in P_0 if condition (5) is satisfied. But Theorem 5 of Ref. 2 implies that condition (5) is satisfied if A is in P_0 and the range of the nonlinearities $f_i(\cdot)$ is the entire real line.* Uniqueness of the solution of equation (2) is very simply shown in Ref. 2. Reference 3 shows that the basic condition, A in P_0, is satisfied for large classes of transistor networks.

The minimum of a continuously differentiable function with bounded level sets can always be found by a gradient descent algorithm when the gradient has a unique root.[6] No assumption regarding convexity or the behavior of the Hessian matrix is necessary. Clearly, a sufficiently small change in $\mathbf{x}$ in the negative gradient direction will always decrease the potential $Q(\mathbf{x})$ unless $\mathbf{x}$ is already at a minimum. A sequence of iterations of this type, that is,

---

* Recently Sandberg[5] has shown that condition (5) holds without any requirements on the range of the nonlinearities if A is nonsingular as well as in P_0.

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \gamma_k \nabla Q(\mathbf{x}_k), \tag{8}$$

monotonically reduces the potential $Q(\mathbf{x})$ and yields a bounded sequence of points $\mathbf{x}_k$ because the level sets are bounded. Convergence of the algorithm (8) is assured if the step sizes can be made large enough so that the potential $Q(\mathbf{x}_k)$ approaches zero rather than a positive limit. This can be achieved by making γ_k depend on the size of the gradient in such a way that γ_k cannot approach zero unless the gradient is approaching zero. Goldstein[6] gives the following procedure for selecting γ_k. Define the normalized potential drop:

$$g(\mathbf{x}, \gamma) = \frac{Q(\mathbf{x}) - Q[\mathbf{x} - \gamma \nabla Q(\mathbf{x})]}{\gamma \| \nabla Q(\mathbf{x}) \|^2}, \qquad \gamma > 0, \tag{9}$$

a continuous function of γ which assumes all values between 1 and 0 as γ ranges between zero and some positive value. Then for any δ with

$$0 < \delta < \tfrac{1}{2}$$

choose γ_k so that

$$\delta \leqq g(\mathbf{x}_k, \gamma_k) \leqq 1 - \delta \tag{10}$$

if $g(\mathbf{x}_k, \gamma_k) < \delta$; otherwise let $\gamma_k = 1$. Note that γ_k can be chosen by trial and error computation in each iteration. For small δ few trials are necessary; but the resulting drop in potential in each iteration is smaller so that more iterations are needed. With this method of choosing γ_k, convergence of the algorithm (8) is assured for any starting point $\mathbf{x}_0$.

In summary, using the optimization approach and a result of Ref. 2 we have shown the existence of a solution to equation (2) and the availability of a convergent algorithm to find the solution under the following conditions.

(I) the nonlinearities $f_i(\cdot)$ are continuously differentiable, strictly monotone increasing, and map the whole real line onto itself, and

(II) the matrix A is in the class P_0.

The original existence conditions given in Ref. 2 do not include the "continuously differentiable" assumption but are otherwise identical to conditions I and II above.

REFERENCES

1. Willson, A. N., Jr., "On the Solutions of Equations for Nonlinear Resistive Networks", B.S.T.J., *47*, No. 8 (October 1968), pp. 1755–1773.
2. Sandberg, I. W., and Willson, A. N., Jr., "Some Theorems on Properties of

dc Equations of Nonlinear Networks", B.S.T.J., *48*, No. 1 (January 1969), pp. 1-34.
3. Sandberg, I. W., and Willson, A. N., Jr., "Some Network-Theoretic Properties of Nonlinear dc Transistor Networks", B.S.T.J., *48*, No. 5 (May-June 1969), pp. 1293-1311.
4. Katzenelson, J., and Seitelman, L. H., "On Iterative Method for Solution of Networks of Nonlinear Monotone Resistors", IEEE Trans. Circuit Theory, *CT-13*, No. 3 (September 1966), pp. 317–323.
5. Sandberg, I. W., unpublished work.
6. Goldstein, A. A., *Constructive Real Analysis*, New York: Harper & Row, 1967, pp. 30–32.